

OPPENHEIM + ZEBRAK, LLP

Matthew J. Oppenheim
Nicholas C. Hailey
Corey Miller
(*pro hac vice* applications to follow)
4530 Wisconsin Ave., NW, 5th Floor
Washington, DC 20016
Telephone: (202) 480-2999
matt@oandzlaw.com
nick@oandzlaw.com
corey@oandzlaw.com

Jennifer Pariser
Carly K. Rothman
Bret Matera
(*pro hac vice* applications to follow)
461 5th Avenue, 19th Floor
New York, NY 10017
Telephone: (212) 951-1156
jpariser@oandzlaw.com
carly@oandzlaw.com
bmatera@oandzlaw.com

Attorneys for Plaintiffs

PRYOR CASHMAN LLP

Benjamin S. Akley (SBN 278506)
1901 Avenue of the Stars, Suite 900
Los Angeles, California 90067
Telephone: (310) 683-6641
bakley@pryorcashman.com

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN JOSE DIVISION**

SONY MUSIC PUBLISHING (US) LLC,
WARNER CHAPPELL MUSIC, INC., ET
AL.,

Plaintiffs,

v.

ANTHROPIC PBC, DARIO AMODEI, AND
BENJAMIN MANN,

Defendants.

**COMPLAINT AND DEMAND FOR
JURY TRIAL**

Plaintiffs Sony Music Publishing (US) LLC, Colgems-EMI Music Inc., Combine Music Corp., EMI April Music Inc., EMI Blackwood Music Inc., EMI Consortium Music Publishing, Inc., EMI Consortium Songs, Inc., EMI Entertainment World Inc. d/b/a EMI Pop Music Publishing, EMI Feist Catalog Inc., EMI Gold Horizon Music Corp., EMI Robbins Catalog Inc., EMI U

Catalog, Inc., EMI UNART Catalog, Inc., Entertainment World Music Inc., Famous Music LLC, Hipgnosis SC I (Delaware) L.P., Hipgnosis SC II (Delaware) L.P., Hipgnosis Songs Group LLC, Jobete Music Co., Inc., Lyra SA (Delaware) L.P., Lyra SC III (Delaware) L.P., Screen Gems-EMI Music Inc., SMG-Three LLC, Stone Diamond Music Corp. (collectively, “Sony Music Publishing” or “SMP Publishers”), Warner Chappell Music, Inc., Cotillion Music, Inc., Gene Autry’s Western Music Publishing Co., Intersong U.S.A., Inc., Ridgeway Music Co., Inc., Rightsong Music Inc., Unichappell Music Inc., W Chappell Music Corp. d/b/a WC Music Corp., W.C.M. Music Corp., Walden Music Inc., and Warner-Tamerlane Publishing Corp. (collectively, “Warner Chappell Music” or “WCM Publishers,” and together with the “SMP Publishers,” the “Music Publishers”), for their complaint against Defendants Anthropic PBC (“Anthropic”), Dario Amodei, and Benjamin Mann (collectively with Anthropic, “Defendants”), allege, on personal knowledge as to matters relating to themselves, and on information and belief as to all other matters, as set forth below.

NATURE OF THE CASE

1. Plaintiff Music Publishers, a group of the world’s leading music publishers, bring this action to hold accountable the culprits behind one of the largest and most blatant ongoing thefts of intellectual property in history. Defendants Anthropic and its founders Dario Amodei and Benjamin Mann have conducted a brazen campaign of illegally torrenting, scraping, and downloading copyrighted works on a massive scale in order to develop, operate, and reap enormous profits from Anthropic’s “Claude” series of artificial intelligence (“AI”) models. Among the innumerable copyrighted works Defendants illegally harvested to fuel Claude are thousands upon thousands of Music Publishers’ copyrighted musical compositions, including such beloved songs as “Ain’t No Mountain High Enough,” “All I Want for Christmas is You,” “Eye of the Tiger,” “Here Comes Santa Claus,” and “Paper Rings.” In blatant violation of copyright law, Defendants have unlawfully acquired troves of Music Publishers’ musical compositions, and then systematically copied those works multiple times, including as the inputs to train Anthropic’s Claude AI models and in the outputs those models generate. These acts separately and together inflict immense harm on Music Publishers and the songwriters they proudly represent.

2. Defendants can no longer hide their extraordinary theft, and their mass infringement is now well-documented. Another court in this District recently described Anthropic’s actions as “straightforward piracy but at massive scale.” *Bartz v. Anthropic PBC*, 791 F. Supp. 3d 1038, 1064 (N.D. Cal. 2025) (“*Bartz I*”). After that court found Anthropic had illegally torrented over seven million copyrighted books from the notorious online pirate websites known as Library Genesis (“LibGen”) and Pirate Library Mirror (“PiLiMi”), Anthropic settled that copyright infringement class action for \$1.5 billion. But Anthropic clearly considers that to be just the cost of doing business given that its entire business model continues to be built on copyright theft. And \$1.5 billion is obviously not a large enough settlement to deter infringing conduct by a company that has parlayed such mass infringement into a staggering \$2-trillion-dollar valuation.¹ Regardless, Anthropic continues to flagrantly violate Music Publishers’ copyrights in numerous, distinct ways.

3. First, Defendants have engaged in flagrant piracy of Music Publishers’ musical compositions via a mass campaign of illegal torrenting. In June 2021, Mr. Mann, on behalf of Anthropic, and at Dr. Amodei’s direction, “used the infamous BitTorrent protocol” to illegally torrent at least five million copies of pirated books from LibGen. *Bartz I*, 791 F. Supp. 3d at 1046. Likewise, in July 2022, Defendants illegally torrented at least another two million copies of pirated books from PiLiMi.

4. Among the many millions of books that Defendants torrented from these illegal pirate websites were books containing the lyrics and sheet music to hundreds or more of Music Publishers’ copyrighted musical compositions, identified in Exhibit A. Those works include “Livin’ On a Prayer,” “September,” “Great Balls of Fire,” “Ramblin’ Man,” and “Hallelujah.”

5. Dr. Amodei and Mr. Mann are personally liable for their respective roles in this illegal torrenting of pirated copies of Music Publishers’ works from LibGen and PiLiMi. Mr. Mann personally used BitTorrent to unlawfully download and upload via torrenting millions of pirated

¹ See, e.g., Jon Markman, *Anthropic Eyes \$2 Trillion In October IPO, A Record-Breaking Debut*, FORBES (Aug. 13, 2026), <https://www.forbes.com/sites/jonmarkman/2026/08/13/anthropic-eyes-2-trillion-in-october-ipo-a-record-breaking-debut/>.

books from LibGen, including many that contain Music Publishers’ copyrighted musical compositions, and he expressly directed other Anthropic employees to do the same with respect to PiLiMi. Dr. Amodei expressly authorized and directed Mr. Mann’s infringements and those of other Anthropic employees. Without Dr. Amodei’s and Mr. Mann’s direction and approval, none of the foregoing torrenting activity would have occurred. At all relevant times, Dr. Amodei and Mr. Mann knew that the millions of books they torrented were pirated and that torrenting those books from LibGen and PiLiMi was illegal.

6. Music Publishers do not yet know the full extent of Defendants’ unlawful torrenting of their works, nor all the ways in which Defendants have exploited the works that they pirated in this manner, because Defendants continue to conceal the critical details of their infringement. At a minimum, Defendants’ use of BitTorrent to download Music Publishers’ works violated Music Publishers’ exclusive right of reproduction and, because BitTorrent operates by simultaneously uploading any copy that a user downloads, Defendants separately violated Music Publishers’ exclusive right of distribution of those works. Defendants have admitted that they torrented these pirated books in order to amass a vast central library of written texts that Anthropic planned to maintain “forever.” Discovery will reveal the full extent to which Defendants exploited this library of pirated works, including Music Publishers’ works, to train and deploy AI models for commercial purposes and other unlawful ends.

7. Regardless, Defendants’ unauthorized torrenting of Music Publishers’ works is legally indefensible (under the fair use doctrine or otherwise), and each instance in which Defendants torrented one of Music Publishers’ works constitutes a separate and standalone act of infringement, causing extensive harm. Indeed, each pirated work that Defendants torrented was likely shared thousands if not tens of thousands of times, depriving Music Publishers of substantial revenue. Defendants also contributed to the continued viability of BitTorrent and pirate libraries as tools for infringement that only exist as long as they have users.

8. In addition to its mass illegal torrenting, Anthropic has acquired unlawful copies of Music Publishers’ works in a range of other ways. Anthropic has illegally harvested unauthorized copies of Music Publishers’ works by “crawling” and “scraping” Music Publishers’ lyrics from the

internet, including from authorized lyric-display websites like MusixMatch and LyricFind; by “destructively scanning” millions of second-hand copies of books, including books containing the lyrics and sheet music to Music Publishers’ works; and by downloading the contents of other notorious third-party datasets like “Common Crawl,” “The Pile,” and “Books3.” Defendants specifically elected to harvest unauthorized copies of Music Publishers’ works so that they would not have to pay Music Publishers a licensing fee to use their works in AI training and development. Indeed, as already determined by the *Bartz* court, Mr. Mann specifically downloaded copies of pirated books containing Music Publishers’ works from the Books3 library on behalf of Anthropic “to avoid the trouble of paying for them, hoping they might prove useful for training large language models (LLMs) or something else.” *Bartz I*, 791 F. Supp. at 1046.

9. Anthropic then exploited these and other illegal copies of Music Publishers’ works as the input to train its Claude AI models. As discovery will further reveal, Anthropic makes additional unauthorized copies of Music Publishers’ lyrics throughout this training process when it “encodes” Music Publishers’ lyrics into the memory of Claude and when it further “finetunes” Claude’s responses to specifically encourage replication of Music Publishers’ lyrics. The countless works that Anthropic unlawfully copies as training input and during the training process in this manner include the tens of thousands of Music Publishers’ copyrighted musical compositions identified in Exhibit B.

10. As a result of Anthropic’s copying of Music Publishers’ works as the input to train its Claude models, those models have likewise generated copies of Music Publishers’ works via outputs to Anthropic’s users and customers, including Music Publishers’ copyrighted musical compositions identified in Exhibit B. Because Anthropic intentionally trains its Claude models to memorize the lyrics from Music Publishers’ works, these models have been trained to generate verbatim and near-verbatim copies of Music Publishers’ lyrics in response to user queries, as well as AI-produced derivatives of those works, in plain violation of Music Publishers’ rights. Additionally, by training its models with copies of Music Publishers’ lyrics, Anthropic teaches its Claude models how to generate vast quantities of purportedly “new” AI-generated song lyrics, which compete with Music Publishers’ legitimate copyrighted works as harmful market substitutes.

Anthropic's infringing acts inflict substantial harm on Music Publishers in multiple ways, including by depriving Music Publishers of licenses fees, undermining potential and existing licensing markets for Music Publishers' lyrics, and lowering the overall value of Music Publishers' works.

11. Finally, Music Publishers bring this suit to address Anthropic's intentional removal and alteration of Music Publishers' copyright management information ("CMI"), including as to Music Publishers' copyrighted musical compositions identified in Exhibit B, in violation of 17 U.S.C. § 1202. After unlawfully harvesting Music Publishers' works, Anthropic processes the text from those works to remove repetitive "garbage" portions that it does not want to use in training its models and with the knowledge and purpose of concealing its other infringements. But what Anthropic considers "garbage" is actually CMI and vital identifying information protected by federal law. Anthropic's removal of this information denies Music Publishers of valuable attribution, including when its Claude models generate verbatim reproductions of Music Publishers' works without appropriate acknowledgment in response to user prompts.

12. Music Publishers recognize the potential of ethical AI technology, and they have entered licenses permitting the authorized use of their musical compositions in connection with AI. It remains crucial, however, that AI development occurs responsibly and on terms that are agreeable to rightsholders and in a manner that protects their rights, livelihoods, and the creative industries as a whole. Even the most revolutionary of technologies must develop within the bounds of the law, and Anthropic's Claude models are no different.

JURISDICTION AND VENUE

13. This is a civil action in which Music Publishers seek injunctive relief and damages under the Copyright Act, 17 U.S.C. § 101, *et seq.* Accordingly, the Court has original subject matter jurisdiction over Music Publishers' copyright infringement claims pursuant to 28 U.S.C. §§ 1331 and 1338(a).

14. The Court has personal jurisdiction over Anthropic because Anthropic is headquartered in this District, does systematic and continuous business in this District, and has

1 committed acts of infringement within this District, including the reproduction of Music Publishers'
2 copyrighted lyrics and the distribution of Music Publishers' lyrics to California citizens.

3 15. The Court has personal jurisdiction over Dario Amodei and Benjamin Mann
4 because they reside and work in this District and have committed acts of infringement in this
5 District.

6 16. Venue is proper in this District in accordance with 28 U.S.C. §§ 1391(b) and 1400(a)
7 because Anthropic, Dr. Amodei, and Mr. Mann reside or may be found in this District; all
8 Defendants conduct, transact, and solicit business in this District; and because a substantial portion
9 of the events and omissions giving rise to the claims asserted herein occurred in this District.

10 **THE PARTIES**

11 **A. Sony Music Publishing Plaintiffs**

12 17. Plaintiff Sony Music Publishing (US) LLC is a Delaware limited liability company
13 with its principal place of business in New York, New York.

14 18. Plaintiff Colgems-EMI Music Inc. is a Delaware corporation with its principal
15 place of business in New York, New York.

16 19. Plaintiff Combine Music Corp. is a Delaware corporation with its principal place
17 of business in New York, New York.

18 20. Plaintiff EMI April Music Inc. is a Connecticut corporation with its principal place
19 of business in New York, New York.

20 21. Plaintiff EMI Blackwood Music Inc. is a Connecticut corporation with its principal
21 place of business in New York, New York.

22 22. Plaintiff EMI Consortium Music Publishing, Inc. is a New York corporation with
23 its principal place of business in New York, New York.

24 23. Plaintiff EMI Consortium Songs, Inc. is a New York corporation with its principal
25 place of business in New York, New York.

26 24. Plaintiff EMI Entertainment World Inc. is a Delaware corporation with its principal
27 place of business in New York, New York, and does business as EMI Pop Music Publishing.
28

1 25. Plaintiff EMI Gold Horizon Music Corp. is a New York corporation with its
2 principal place of business in New York, New York.

3 26. Plaintiff EMI Feist Catalog Inc. is a New York corporation with its principal place
4 of business in New York, New York.

5 27. Plaintiff EMI Robbins Catalog Inc. is a New York corporation with its principal
6 place of business in New York, New York.

7 28. Plaintiff EMI U Catalog, Inc. is a New York corporation with its principal place of
8 business in New York, New York.

9 29. Plaintiff EMI UNART Catalog, Inc. is a New York corporation with its principal
10 place of business in New York, New York.

11 30. Plaintiff Entertainment World Music Inc. is a Delaware corporation with its
12 principal place of business in New York, New York.

13 31. Plaintiff Famous Music LLC is a Delaware corporation with its principal place of
14 business in New York, New York.

15 32. Plaintiff Hipgnosis Songs Group LLC is a Delaware limited liability corporation
16 with its principal place of business in New York, New York.

17 33. Plaintiff Hipgnosis SC I (Delaware) L.P. is a Delaware limited partnership with its
18 principal place of business in New York, New York.

19 34. Plaintiff Hipgnosis SC II (Delaware) L.P. is a Delaware limited partnership with its
20 principal place of business in New York, New York.

21 35. Plaintiff Jobete Music Co., Inc. is a Michigan corporation with its principal place
22 of business in New York, New York.

23 36. Plaintiff Lyra SA (Delaware) L.P. is a Delaware limited partnership with its
24 principal place of business in New York, New York.

25 37. Plaintiff Lyra SC III (Delaware) L.P. is a Delaware limited partnership with its
26 principal place of business in New York, New York.

27 38. Plaintiff Screen Gems-EMI Music Inc. is a Delaware corporation with its principal
28 place of business in New York, New York.

39. Plaintiff SMG-Three LLC is a Delaware limited liability company with its principal place of business in New York, New York.

40. Plaintiff Stone Diamond Music Corp. is a Michigan corporation with its principal place of business in New York, New York.

B. Warner Chappell Music Plaintiffs

41. Plaintiff Warner Chappell Music, Inc. is a Delaware corporation with its principal place of business in Los Angeles, California, and does business as Chappell & Co. and Warner Geometric Music.

42. Plaintiff Cotillion Music, Inc. is a Delaware corporation with its principal place of business in Los Angeles, California.

43. Plaintiff Gene Autry's Western Music Publishing Co. is a California corporation with its principal place of business in Los Angeles, California.

44. Plaintiff Intersong U.S.A., Inc. is a Delaware corporation with its principal place of business in Los Angeles, California.

45. Plaintiff Ridgeway Music Co., Inc. is a California corporation with its principal place of business in Los Angeles, California.

46. Plaintiff Rightsong Music Inc. is a Delaware corporation with its principal place of business in Los Angeles, California.

47. Plaintiff Unichappell Music Inc. is a Delaware corporation with its principal place of business in Los Angeles, California.

48. Plaintiff W Chappell Music Corp. is a California corporation with its principal place of business in Los Angeles, California, and also does business as WC Music Corp.

49. Plaintiff W.C.M. Music Corp. is a Delaware corporation with its principal place of business in Los Angeles, California.

50. Plaintiff Walden Music Inc. is a New York corporation with its principal place of business in Los Angeles, California.

51. Plaintiff Warner-Tamerlane Publishing Corp. is a California corporation with its principal place of business in Los Angeles, California.

1 **C. Defendant Anthropic**

2 52. Defendant Anthropic PBC is a Delaware corporation with its principal place of
3 business in San Francisco, CA.

4 53. Defendant Dario Amodei is an individual residing in San Anselmo, California. He
5 is a founder and Chief Executive Officer of Anthropic.

6 54. Defendant Benjamin Mann is an individual residing in or around San Francisco,
7 California. He is a co-founder and Member of the Technical Staff of Anthropic.

8 **PLAINTIFF MUSIC PUBLISHERS AND THEIR COPYRIGHTED WORKS**

9 55. Music Publishers are in the business of creating, developing, promoting, publishing,
10 administering, acquiring, and licensing musical compositions. Their catalogs include some of
11 history's most iconic songs, performed by the world's most legendary artists, with lyrics that move
12 and inspire fans of all ages.

13 56. Songwriting is no easy task: it takes talent, creativity, emotion, resources, and more
14 to write memorable and popular lyrics. Music Publishers, and the songwriters they represent, rely
15 upon our nation's copyright laws to protect their musical compositions (and the lyrics therein)
16 from infringement. Copyright fosters Music Publishers' incentive to develop, promote, publish,
17 administer, acquire, license, and otherwise invest in songwriters and their musical compositions.
18 Without copyright, Music Publishers and their songwriters lose control of their own artistic
19 expression, and with it, the incentive to create and the ability to earn a living.

20 57. Music Publishers own and/or control, in whole or in part, the exclusive rights to
21 millions of musical compositions, including those listed in Exhibits A and B, which are illustrative
22 and non-exhaustive lists of the musical compositions infringed by Anthropic.² Music Publishers
23 own or exclusively license the copyrights in and to these works in a variety of ways, including
24 pursuant to an agreement with either the songwriter(s) or a predecessor publisher. All the musical
25 compositions listed in Exhibits A and B have been registered with the U.S. Copyright Office. The
26

27 ² Music Publishers' investigation is continuing, and discovery is likely to reveal additional
28 infringements. Music Publishers intend to seek leave to amend this Complaint at an appropriate
time to provide an expanded list of musical compositions infringed by Defendants.

1 copyrights in the musical compositions set forth in Exhibits A and B remain valid and subsisting
2 and have been owned and/or controlled by Music Publishers at all times relevant to the allegations
3 in this Complaint.

4 58. Music Publishers own and/or exclusively control rights set forth in 17 U.S.C. § 106
5 of the Copyright Act with respect to the musical compositions set forth in Exhibits A and B, and
6 many more. Anyone who wants to use the musical compositions owned and/or controlled by Music
7 Publishers for commercial purposes, as Defendants have done, must comply with copyright law
8 by obtaining necessary licenses and permissions.

9 59. Each of the rights under the Copyright Act, including to reproduce the work,
10 prepare derivative works, and distribute or display copies to the public, is a separate right within
11 the bundle of rights that a copyright owner (or holder of an exclusive license) is entitled to exploit.
12 These exclusive rights are reserved under copyright law for those that own or control the exclusive
13 rights, and Anthropic is not authorized to appropriate them to suit its business model.

14 60. Music Publishers license and otherwise exploit these rights, among others, to earn
15 revenue from their musical compositions. This revenue is typically shared with the songwriters
16 whom Music Publishers represent to ensure they are compensated for the use of their creative
17 efforts. Songwriters depend on this compensation to earn a living while enriching the world with
18 new music. A legal system of licensing, permission, and authorization enables this relationship
19 between songwriters and Music Publishers and ensures all are compensated, while also
20 safeguarding the artistic integrity and control of the musical works.

21 61. Music Publishers embrace technological innovation and provide their songs and
22 catalogs across an ever-evolving variety of formats, distribution platforms, and access models.
23 Today, Music Publishers and their partners offer consumers a variety of authorized means to enjoy
24 Music Publishers' musical compositions, including through licensed print and digital books
25 containing sheet music and lyrics, as well as through licensed lyric aggregators, lyric websites,
26 social media platforms, and a host of other online platforms. Through these and other licensed uses,
27 Music Publishers ensure that consumers can access and appreciate genuine, authorized copies of
28 the music and lyrics to their favorite songs.

62. To that end, Anthropic is wrong to claim that Music Publishers’ lyrics are “freely” available on the internet. Lyric websites pay for the right to display Music Publishers’ lyrics and to commercialize that display with advertising revenue, a portion of which is typically returned to Music Publishers under the terms of their licenses. By licensing display rights to lyric websites, Music Publishers have not authorized AI companies like Anthropic to systematically copy their lyrics from those websites and use them *en masse* to train and develop commercial AI models, thereby profiting off of Music Publishers’ works without authorization or compensation.

63. Additionally, when Music Publishers license their works, such as to authorized lyric aggregators and websites, the licensees are typically required to identify such lyrics with the song title, songwriter name(s), and other important identifying information, all of which constitutes CMI. This information is vital to consumers, who want to know the authors of these musical compositions, and it is vital to songwriters and Music Publishers, who want to ensure that the creators and owners of these musical works are properly credited and that the public is informed that these musical works are protected by copyright. Music Publishers therefore typically require licensees to display CMI, and its removal by unlicensed third parties violates Music Publishers’ rights and harms their interests and those of their songwriters.

64. Music Publishers presently work with multiple AI companies to pioneer innovative and highly successful frameworks for users to make lawful use of Music Publishers’ songs on those platforms. Music Publishers have entered into agreements with AI companies covering use of Music Publishers’ musical compositions on platforms with which Anthropic competes.

ANTHROPIC AND ITS INFRINGEMENT OF MUSIC PUBLISHERS’ WORKS

A. Anthropic’s Business

65. Anthropic develops, operates, sells, and licenses access to generative AI programs. After leaving rival AI company OpenAI, Defendants Dario Amodei and Benjamin Mann co-founded Anthropic in 2021 with a group of other former OpenAI employees. Since then, Anthropic has skyrocketed to a “post-money valuation” of \$965 billion,³ before jumping to an even more

³ *Anthropic raises \$65B in Series H funding at \$965B post-money valuation*, ANTHROPIC (May 29, 2026), <https://www.anthropic.com/news/series-h>.

1 staggering \$2-trillion-dollar valuation,⁴ and is on its way to becoming the most valuable AI start-
2 up company in the world. But despite its multi-trillion-dollar valuation, Anthropic refuses to pay
3 for the vast quantities of copyrighted content, including Music Publishers’ works, that it harvests
4 to train the machines at the heart of its business.

5 66. Anthropic’s signature product is a series of commercial general-purpose large
6 language models (“LLMs”) referred to as “Claude.” Anthropic “trains” these AI models with a
7 massive hoard of text-based data—including the lyrics to Music Publishers’ copyrighted works,
8 totaling billions or trillions of words. From this training, Anthropic’s AI models learn to respond
9 to user prompts, including those relating to song lyrics, with text-based outputs in pseudo-human
10 fashion. Anthropic has never sought nor secured Music Publishers’ permission to use their valuable
11 copyrighted works for this or any other purpose.

12 67. Anthropic collects and stockpiles its harvested data into a vast central “library” that
13 it maintains and exploits for a range of commercial and research purposes, including but not limited
14 to training Claude models. Anthropic compiled this central library by copying and ingesting text
15 from a variety of sources (many of them well-known to be illegal) and via a range of methods,
16 including by using BitTorrent to download and copy text from illegal pirate websites. Through
17 these sources and methods, Anthropic unlawfully copied and ingested the lyrics to an untold
18 number of Music Publishers’ copyrighted musical compositions into the company’s central library.

19 68. Anthropic began developing and training its AI models in January 2021 and
20 released its signature Claude models to the general public beginning in March 2023. Anthropic
21 regularly releases newer and more powerful Claude models every few months. Anthropic trains
22 each of its AI models, including all public versions of Claude, with unauthorized copies of Music
23 Publishers’ lyrics and other copyrighted works. Each Claude model is trained using a newly copied
24 “training corpus” that contains Music Publishers’ copyrighted works.

25 69. Anthropic provides access to its Claude AI models in multiple ways, including
26

27 ⁴ See, e.g., Jon Markman, *Anthropic Eyes \$2 Trillion In October IPO, A Record-Breaking Debut*,
28 FORBES (Aug. 13, 2026), <https://www.forbes.com/sites/jonmarkman/2026/08/13/anthropic-eyes-2-trillion-in-october-ipo-a-record-breaking-debut/>.

1 through its Claude.ai website, via mobile apps, and as a commercial application programming
2 interface (“API”) through which custom third-party client software can interact with Claude AI
3 models. Claude responds to user prompts and commands with conversational, pseudo-human
4 replies. More sophisticated versions of Claude are available via paid subscription. Anthropic sells
5 or licenses Claude API access to commercial customers, which includes incorporation of Claude
6 into these customers’ own proprietary software, products, and systems. When a commercial
7 customer prompts Claude through their own software using the Anthropic API, the customer’s
8 software sends the prompt to an Anthropic service, at which point the prompt passes as input into
9 the Claude model. Anthropic then sends the model’s response back to the API customer’s software.

10 70. To date, Anthropic has publicly released the following commercial Claude models,
11 all of which have been trained on unauthorized copies of Music Publishers’ works: Claude 1 (Mar.
12 2023), Claude 2 (July 2023), Claude 2.1 (Nov. 2023), Claude 3 Sonnet and Claude 3 Opus (Mar.
13 2024), Claude 3.5 (June 2024), Claude 3.7 Sonnet (Feb. 2025), Claude 4.5 Sonnet (Sept. 2025),
14 Claude 4.5 Haiku (Oct. 2025), Claude 4.5 Opus (Nov. 2025), Claude 4.6 Opus (Feb. 2026), Claude
15 4.6 Sonnet (Feb. 2026), Claude Opus 4.8 (May 2026), Claude Sonnet 5 (June 2026), Claude Fable
16 5 (June 2026), and Claude Opus 5 (July 2026). Anthropic also used Music Publishers’ works to
17 train numerous non-public AI models (for commercial and research purposes, or otherwise) and is
18 currently using Music Publishers’ works to train versions of Claude and other AI models that it has
19 not yet released to the public.

20 71. Despite branding itself as the “ethical AI company,” Anthropic has repeatedly acted
21 in ways that belie that image, prioritizing competitive advantage over compliance with the law.
22 Indeed, while Anthropic was originally founded by a group of former OpenAI employees
23 purportedly intent on building AI models with a focus on safety, earlier this year the company
24 abandoned the core safety commitment that it first made in 2023: to stop training and refuse to
25 deploy an AI system if it could not guarantee it had proper safety guardrails in place that met
26 stringent internal standards.⁵

27
28 ⁵ See *Responsible Scaling Policy Version 3.0*, ANTHROPIC (Feb. 24, 2026), <https://www-cdn.anthropic.com/e670587677525f28df69b59e5fb4c22cc5461a17.pdf>.

72. In an interview with *Time* regarding Anthropic’s decision to drop its core safety pledge, Anthropic’s founder and Chief Science Officer Jared Kaplan succinctly described Anthropic’s true ethos: “We didn’t really feel, with the rapid advance of AI, that it made sense for us to make unilateral commitments . . . if competitors are blazing ahead.”⁶ An Anthropic researcher recently echoed this same sentiment in his open exit letter from the company, where he observed that “[t]hroughout my time [at Anthropic], I’ve repeatedly seen how hard it is to truly let our values govern our actions” and that employees “constantly face pressure to set aside what matters most.”⁷ What’s more, Anthropic was recently forced to disclose that three of its AI models—Claude Opus 4.7, Claude Mythos 5, and an internal research test model—independently broke out of their testing environments, accessed the internet, and went on to hack three different companies, “compromis[ing] the impacted organizations’ infrastructure.”⁸

B. Anthropic’s Infringement via Torrenting

73. From its inception, Anthropic faced stiff competition on the AI frontier from some of the world’s most powerful tech companies, including Google, Microsoft, Meta, and newcomer OpenAI. If it was going to gain market share, Anthropic knew that it needed to get its Claude models off the ground—and fast—which meant that it needed as much high-quality training data that it could get its hands on. To do so, Anthropic torrented millions of books from two so-called “pirate libraries,” notorious for containing vast collections of illicitly obtained content: LibGen and PiLiMi.

74. Defendants’ torrenting conduct began in June 2021, when Mr. Mann personally used “the infamous BitTorrent protocol” to unlawfully torrent via BitTorrent at least five million books from LibGen. *Bartz I*, 791 F. Supp. at 1046.

⁶ Billy Perrigo, *Exclusive: Anthropic Drops Flagship Safety Pledge*, TIME (Feb. 24, 2026), <https://time.com/7380854/exclusive-anthropic-drops-flagship-safety-pledge/>.

⁷ Frank Landymore, *Anthropic Researcher Quits in Cryptic Public Letter*, FUTURISM (Feb. 10, 2026), <https://futurism.com/artificial-intelligence/anthropic-researcher-quits-cryptic-letter>.

⁸ *Investigating three real-world incidents in our cybersecurity evaluations*, ANTHROPIC (July 30, 2026), <https://www.anthropic.com/news/investigating-incidents-cybersecurity-evals>.

1 75. LibGen is one of the world’s largest and most notorious online pirate libraries.
2 Anonymous overseas operators, reportedly based in Russia, use it to host and distribute an
3 enormous collection of stolen books and other written works. At least two U.S. federal courts have
4 found LibGen liable for willful copyright infringement and issued broad permanent injunctions
5 against the website. LibGen has also appeared on the U.S. Trade Representative’s “Review of
6 Notorious Markets for Counterfeiting and Piracy” since at least 2016, and its unlawful activities
7 are widely known and well-documented.

8 76. BitTorrent is a peer-to-peer (“P2P”) filesharing protocol most commonly used for
9 widespread unauthorized reproduction and distribution of copyrighted materials. BitTorrent allows
10 a user to break data files into pieces, which are then distributed between and among multiple users
11 (or “peers”), simultaneously and instantaneously. When a peer downloads a piece of a file, that
12 peer immediately begins sharing that piece with other peers, creating what is known as a “swarm,”
13 where everyone downloading the file simultaneously acts as a distributor of that file to others.
14 Swarms are an effective and fast means of distribution because they can potentially include
15 thousands of simultaneous peers. And once an infringing file is distributed on BitTorrent, there is
16 no putting the genie back in the bottle. The viral nature of online infringement enables massive
17 rapid proliferation across a limitless number of peers in mere hours.

18 77. With Dr. Amodei’s knowledge, approval, and participation, Mr. Mann used
19 BitTorrent to torrent more than five million books from LibGen, including at least hundreds of
20 songbooks and sheet music containing the lyrics to hundreds or more of Music Publishers’
21 copyrighted works. Mr. Mann discussed his torrenting in real time on Anthropic’s internal Slack
22 channels and even shared a screenshot of his torrenting activity with Anthropic colleagues. To
23 address the concern that the sheer size of the LibGen torrent would exhaust his computer storage,
24 he personally set up a computer program to “pause torrenting if disk space got ‘exhausted,’” *Bartz*
25 *I*, 791 F. Supp. 3d at 1046, which he flippantly described as “a cute little libgen babysitter,” *Bartz*
26 *et al. v. Anthropic PBC*, 4:24-cv-05417-AMO (N.D. Cal.) (hereinafter, “*Bartz Docket*”), ECF No.
27 560-30 at 4.

28 78. This was not Mr. Mann’s first time torrenting pirated content from LibGen either.

1 Before co-founding Anthropic, he personally downloaded a portion of LibGen to be used in AI
2 training at OpenAI, which he discussed with both Dr. Amodei and another future Anthropic
3 founder, Tom Brown. Mr. Mann later testified that he discussed the legality of torrenting LibGen
4 with Anthropic's attorneys but would not say under oath whether it was "exactly lawful or not."
5 *Bartz* Docket, ECF No. 337-1 at 89:18–90:5, 96:7-22. Public documents reveal, however, that Mr.
6 Mann was well aware of LibGen's reputation as a source of pirated content even before joining
7 Anthropic, characterizing the pirate website in internal Anthropic discussions as "sketchy AF."
8 *Bartz* Docket, ECF No. 325 at 8.

9 79. Mr. Mann also consulted Dr. Amodei and another Anthropic co-founder, Chief
10 Science Officer Jared Kaplan, before torrenting LibGen for Anthropic. At the time, Dr. Amodei
11 worked closely on the development of Anthropic's AI models, and Anthropic's policy required his
12 express approval before Mr. Mann (or others) could obtain or use new datasets in AI model
13 development. To that end, pursuant to Anthropic's own policy, Mr. Mann required Dr. Amodei's
14 express approval before torrenting LibGen, notwithstanding the fact that Anthropic's own Archive
15 Team described LibGen as a "blatant violation of copyright." *Bartz* Docket, ECF No. 337-1 at
16 144:4-13, 396:3-13. Indeed, Mr. Mann admitted in a recent court pleading that "he discussed the
17 potential acquisition of LibGen data with Dr. Amodei and other Anthropic colleagues," and that
18 "Dr. Amodei approved" his torrenting of LibGen. *Concord Music Group, Inc. et al v. Anthropic*
19 *PBC et al.*, 5:26-cv-00880-EKL (N.D. Cal.) ("*Concord II* Docket"), ECF No. 106 ("Answer to
20 Amended Complaint by Benjamin Mann") at ¶ 71. Accordingly, at the direction of Dr. Amodei,
21 Mr. Mann proceeded to personally torrent millions of books from LibGen. Dr. Amodei and Mr.
22 Mann knew and understood that LibGen was an illegal pirate website but proceeded to torrent its
23 contents anyway, because doing so was cheaper than "buy[ing] all the books in libgen." *Bartz I*,
24 791 F. Supp. 3d at 1047; *see also Bartz* Docket, ECF No. 563-5 at 2. Indeed, the *Bartz* court
25 previously determined as fact that Mr. Mann knew that the five million copies of books that he
26 torrented from LibGen "had been pirated." *Bartz v. Anthropic PBC*, 787 F. Supp. 3d 1007, 1015
27 (N.D. Cal. 2025) ("*Bartz II*").
28

80. In late 2021, the FBI and the U.S. government shut down LibGen, but not before online pirates copied its contents to create a new library named “Z-Library.” The FBI quickly shut down Z-Library as well, but not before it was copied, or “mirrored,” yet again in July 2022, this time under the name “Pirate Library Mirror,” or “PiLiMi.” *Id.* The operators of PiLiMi declare openly that they “deliberately violate the copyright law in most countries.”⁹

81. When Mr. Mann learned that PiLiMi was ready to be torrented, he shared the link to PiLiMi and messaged his Anthropic colleagues, “[J]ust in time!” *Bartz I*, 791 F. Supp. 3d at 1046. Another Anthropic employee responded, “zlibrary my beloved.” *Id.* The common lineage between PiLiMi and LibGen through Z-Library is significant because both libraries contain many of the same pirated books. Appreciating this overlap, Anthropic engineers compared the five million copies that Mr. Mann had previously torrented from LibGen against seven million copies available to torrent from PiLiMi and proceeded to torrent the two million copies that were not included in LibGen. *Id.*

82. Accordingly, in July 2022, certain Anthropic employees, acting at Mr. Mann’s and Dr. Amodei’s direction, and with their approval, torrented millions of additional copies of pirated books, including books containing Music Publishers’ copyrighted musical compositions, from distributed copies of PiLiMi. Dr. Amodei, Mr. Mann, and their Anthropic colleagues understood that the PiLiMi files were pirated and sourced from “a popular (and illegal) library,” *Bartz Docket*, ECF No. 560-25 at 2, but authorized the torrenting anyway.

83. In total, Defendants torrented at least five million books from LibGen and at least two million books from PiLiMi. *Bartz I*, 791 F. Supp. 3d at 1046. Non-confidential catalogs of LibGen and PiLiMi, reflecting bibliographic metadata like title, author, and ISBN, reveal that Anthropic torrented at least hundreds of books containing sheet music and song lyrics to musical compositions owned by Music Publishers, as well as standalone copies of sheet music and other texts containing Music Publishers’ musical compositions, including but not limited to the works

⁹ Ernesto Van der Sar, “*Anna’s Archive*” Opens the Door to Z-Library and Other Pirate Libraries, TORRENTFREAK (Nov. 19, 2022), <https://torrentfreak.com/annas-archive-opens-the-door-to-z-library-and-other-pirate-libraries-221118/>.

1 identified in Exhibit A.

2 84. Defendants claim to have torrented these millions of copyrighted works to build a
3 vast central library for Anthropic’s unrestricted use, including for the development of its
4 commercial Claude AI products. *Bartz I*, 791 F. Supp. 3d at 1047. Defendants continue to maintain
5 and store copies of these torrented files in their original formats, including as full-text digital
6 books—“ebooks or scans of books” saved in individual files in formats like .pdf, .txt, and .epub.
7 *Bartz II*, 787 F. Supp. 3d at 1015. Anthropic has stated that it intends to store copies of all
8 copyrighted works it has illegally harvested, including Music Publishers’ musical compositions,
9 “forever.” *Id.* at 1016.

10 85. Defendants publicly deny using any of the books that they torrented from LibGen
11 and PiLiMi to train Anthropic’s commercial Claude AI models. But that denial is based on how
12 Anthropic defines “training.” Often, AI development includes a “pretraining” phase where one AI
13 model may train on “synthetic data” created by another AI model or receive other behavioral
14 feedback from another AI model. Anthropic has trained at least one of its commercially released
15 Claude models using synthetic data created by a non-commercial AI model that was trained on
16 text derived from LibGen and/or PiLiMi. Anthropic employed at least one non-commercial model
17 that was trained on text derived from LibGen and/or PiLiMi to provide at least one commercial
18 Claude model with reinforced feedback. And discovery will reveal the full extent to which
19 Anthropic utilized the books it pirated from LibGen and PiLiMi, including Music Publishers’
20 works, to train its Claude models more broadly.

21 86. Regardless of subsequent use, Anthropic’s torrenting was unquestionably
22 infringing. To the extent Anthropic made additional copies of any illegally torrented books in order
23 to exploit those works for any purpose (research, commercial, or otherwise), the act of making
24 such copies constitutes an additional infringing act. Moreover, Anthropic maintained copies of
25 these illegally torrented files as part of its central library.

26 87. Between LibGen and PiLiMi, Defendants illegally torrented over seven million
27 pirated books. Defendants’ illicit haul included large numbers of books from Music Publishers’
28 sheet music licensees. Those titles include *The Beatles Complete Scores: Every Song Written &*

1 *Recorded By the Beatles* (featuring “Twist and Shout” and other works); *Best of Taylor Swift*
2 *Songbook* (featuring “Cruel Summer” and other works); *Bon Jovi These Days* (featuring “These
3 Days” and other works), among many others. These books contain properly licensed copies of
4 Music Publishers’ sheet music and song lyrics and CMI attributed to each work. These books also
5 typically state on their face that unauthorized copying of the works contained therein constitutes
6 copyright infringement—a warning Defendants blatantly ignored.

7 88. When Defendants torrented these books, they violated Music Publishers’ exclusive
8 right of reproduction in these works. For those pirated works that Defendants downloaded via
9 torrenting, they also simultaneously uploaded unauthorized copies of the same books due to the
10 two-way nature of the BitTorrent protocol, infringing Music Publishers’ exclusive right of
11 distribution in the process and contributing to further infringement of Music Publishers’ works.

12 89. Defendants knowingly exploited LibGen and PiLiMi to remain competitive with
13 OpenAI, as well as other established technology companies with significant data advantages, like
14 Google, Meta, and Microsoft. Defendants sought to amass a library of “all the books in the world”
15 to retain “forever,” *Bartz II*, 787 F. Supp. 3d at 1014, and to do so as quickly—and cheaply—as
16 possible. Dr. Amodei admitted that Anthropic “ha[d] many places from which” it could have
17 legally purchased these copyrighted works for training, but instead torrented them because doing
18 so was faster and free. *Id.* at 1015. In Dr. Amodei’s own words, they did so to avoid a
19 “legal/practice/business slog.” *Id.* As the *Bartz* court put it, Anthropic “downloaded the pirated
20 books to avoid the trouble of paying for them.” *Bartz I*, 791 F. Supp. 3d at 1046.

21 90. Defendants knew that these pirate libraries were not legal sources of copyrighted
22 works. One unsealed Anthropic memo expressly notes that “books published after 1929 found on
23 the Web” are often “pirated and posted illegally.” *Bartz* Docket, ECF No. 554-29 at 6–7.
24 Defendants made the deliberate choice to prioritize Anthropic’s competitive and financial interests
25 over copyright holders’ rights and then took multiple steps to conceal their infringing activity. Even
26 after Anthropic became “not so gung ho about” training AI models using pirated books “for legal
27 reasons,” *Bartz II*, 787 F. Supp. 3d at 1015, it maintained copies of its pirated files in its central
28 library anyway, ignoring that they are copyrighted works that it obtained illegally.

C. **Anthropic’s Infringement via Scraping, Downloading, and Destructive Scanning**

91. Anthropic’s practice of illegal data harvesting does not stop with its exploitation of pirate libraries and BitTorrent. With the assistance of its own “web crawlers,” Anthropic also copies—or “scrapes”—enormous volumes of copyrighted text from the internet without authorization. Anthropic also uses datasets that it has obtained from third parties and elsewhere—comprised of scraped content, pirated content, and content from various other sources—all without permission from the appropriate rightsholders. Anthropic copies and downloads vast amounts of text and raw data copies from these sources onto its own servers, which it subsequently curates for AI training.

92. Among the billions—if not trillions—of texts that Anthropic copies from online sources are innumerable copies of Music Publishers’ copyrighted works. Anthropic scraped and copied these works itself, without Music Publishers’ permission, including from websites expressly authorized to display their works, like MusixMatch.com and LyricFind.com, in direct violation of those websites’ terms and conditions. Discovery will reveal the full scope of Anthropic’s copying of Music Publishers’ works, which facts are in the possession and control of Anthropic.

93. Limited public disclosures have confirmed that Anthropic has also used and exploited numerous third-party datasets known to contain unauthorized copies of Music Publishers’ lyrics and musical compositions. For example, the *Bartz* court found that, mere months after Anthropic’s founding in early 2021, Mr. Mann “downloaded 196,640 unauthorized copies of copyrighted books” from the pirate dataset known as “Books3.” *Bartz I*, 791 F. Supp. 3d at 1046. Anthropic has admitted to using Books3 for AI training with the knowledge that Books3 contained at least “183k pirated books,” with one internal memo endorsing the pirate dataset for its “high quality” while also noting its highly questionable “legal status.” *Bartz* Docket, ECF No. 554-28 at 2–3.

94. Books3 was created by an independent developer named Shawn Presser, who compiled the dataset with pirated books that he downloaded from a website called Bibliotik in an

attempt to “reverse-engineer” two datasets that OpenAI used to train its early GPT-3 model: “books1” and “books2.” *Bartz* Docket, ECF 560-45 at 3. Books1 and books2 were predecessor collections of thousands of pirated books, torrented from a variety of illegal websites—and compiled by none other than Dr. Amodei and Mr. Mann while they were still employed at OpenAI.¹⁰ OpenAI has since destroyed “books1” and “books2,” but their legacy lives on at Anthropic through the pirated contents of Books3, which, as the *Bartz* court described, Mr. Mann downloaded “to avoid the trouble of paying for [it], hoping [the contents] might prove useful for training large language models (LLMs) or for something else.” *Bartz I*, 791 F. Supp. 3d at 1046.

95. Anthropic has also admitted to copying and exploiting for use in AI training a well-known dataset called “The Pile.”¹¹ The Pile incorporates several preexisting “high-quality datasets,”¹² including another full copy of the illegal “Books3” dataset, as well as a “YouTube Subtitles” dataset, which consists of the “human generated closed captions” for 173,651 YouTube videos,¹³ including official music videos and lyric videos for compositions owned and controlled by Music Publishers. Anthropic has admitted to using “YouTube transcripts from the Pile” to train its AI models even though an internal Anthropic memo noted the questionable “legal status” of “[t]ranscribing the audio from scraped Youtube videos.” Anthropic continues to employ The Pile to train its most recent Claude models. *Bartz* Docket, ECF No. 554-28 at 4.

96. In addition to harboring pirated and unauthorized copies of copyrighted text, Defendants understood that The Pile’s creators also removed copyright notices and ownership information. In a widely published industry whitepaper, The Pile’s creators publicly disclosed that

¹⁰ See *Bartz* Docket, ECF 337-1 at 94:5-96:6, 349:21-23; see also *In re Motion to Compel Compliance with Non-party Subpoenas*, No. 25-mc-80017, ECF No. 17 at 1–3 (N.D. Cal. Feb. 24, 2025) (identifying Dr. Amodei and Mr. Mann as creators of “books1” and “books2”).

¹¹ Amanda Askeel et al., *A General Language Assistant as a Laboratory for Alignment*, 27, ANTHROPIC (Dec. 9, 2021), <https://arxiv.org/pdf/2112.00861>; see Samuel Axon, *YouTube Creators Surprised to Find Apple and Others Trained AI on Their Videos*, ARS TECHNICA (July 16, 2024), <https://arstechnica.com/ai/2024/07/apple-was-among-the-companies-that-trained-its-ai-on-youtube-videos>.

¹² Leo Gao et al., *The Pile: An 800GB Dataset of Diverse Text for Language Modeling*, 2, ARXIV (Dec. 31, 2020), <https://arxiv.org/pdf/2101.00027>.

¹³ *Id.* at 5.

1 the YouTube Subtitles dataset was “not collected and distributed in a ToS [terms of service]
2 compliant fashion” and that “the authors had no ability to consent to their data being used.”¹⁴
3 Similarly, The Pile creators disclosed their use of extraction tools, like Newspaper and Dragnet,¹⁵
4 which are commonly used within the AI community, including by Anthropic, to remove undesired
5 CMI from text content scraped from a website. The developers of Claude cited this paper and knew
6 that these tools were used to prune The Pile. Dragnet’s algorithms, in particular, separate main
7 website content, such as lyrics, from footers, copyright notices, and other CMI, allowing extraction
8 of only the main content.

9 97. Anthropic also copied another notorious dataset known as “Common Crawl,”
10 which Anthropic continues to exploit in connection with ongoing AI training. The earliest version
11 of Common Crawl was compiled in 2013, long before the advent of generative AI, and likewise
12 contains innumerable copies of Music Publishers’ copyrighted lyrics, scraped without permission
13 from the websites of Music Publishers’ licensees and other entities that Music Publishers have
14 authorized to reproduce their copyrighted lyrics, such as MusixMatch and LyricFind.¹⁶

15 98. Anthropic’s data harvesting is not limited to the digital world. Anthropic employs
16 a sophisticated “destructive scanning” operation, modeled after the Google Books project, which
17 involves purchasing physical copies of millions of second-hand books, scanning ink and paper into
18 digital text, then destroying the remains. Once digitized, Anthropic adds the newly scanned text to
19 its central library for potential later use in AI training. As with the millions of pirated books that
20 Defendants torrented from LibGen and PiLiMi, Anthropic’s destructive scanning project has
21 resulted in the unauthorized digitization of hundreds of songbooks and sheet music collections
22 containing the copyrighted musical compositions owned and controlled by Music Publishers and
23 others, in violation of their copyrights.

24 99. But unlike Google Books, Anthropic does not provide the public with a free,

25 ¹⁴ *Id.* at 14.

26 ¹⁵ *See id.* at 22–23.

27 ¹⁶ Kevin Schaul et al., *Inside the Secret List of Websites That Make AI Like ChatGPT Sound*
28 *Smart*, WASHINGTON POST (Apr. 19, 2023),
<https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning>.

1 searchable database to help readers discover new books. Rather, Anthropic’s goal is to acquire “all
2 the books in the world” to power its own enterprise in a way that avoids as much
3 “legal/practice/business slog” as possible. *Bartz II*, 787 F. Supp. 3d at 1014–15. But even the
4 company’s own employees recognize that this destructive practice is emblematic of how many
5 perceive AI to be destroying the arts. Indeed, a recently unsealed planning document from 2024
6 stated, “We don’t want it to be known that we are working on this.” *Bartz* Docket, ECF No. 554-
7 21 at 2.

8 100. Discovery will reveal the full extent to which Anthropic has infringed Music
9 Publishers’ copyrighted works and included them in Anthropic’s training data and central library.

10 **D. Anthropic’s Infringement via Removal of Copyright Management**
11 **Information**

12 101. Anthropic employs an army of dedicated engineers tasked with curating subsets of
13 specific texts (and combinations thereof) that comprise the “training corpus” for each new iteration
14 of Claude. During this curation process, Anthropic engages in additional acts of unauthorized
15 copying and infringement.

16 102. Most notably, Anthropic violates Music Publishers’ rights when it “cleans” the text
17 it harvests to remove material it perceives as undesirable for AI training. To be clear, this “cleaning”
18 process does not remove unauthorized copyrighted content, such as Music Publishers’ lyrics, from
19 the training corpus. Rather, the “cleaning” process removes corresponding CMI embodied in the
20 copied text, including in Music Publishers’ lyrics, and other attribution and ownership information
21 necessary to identify these copies as infringements. Anthropic takes these steps to conceal its
22 underlying infringement of Music Publishers’ works.

23 103. Anthropic knows that, after training on Music Publishers’ works, Claude is likely
24 to generate output containing Music Publishers’ CMI unless Anthropic removes that information
25 from its AI training corpus. Accordingly, Anthropic intentionally extracts CMI from its training
26 data in order to prevent its models from overweighting and displaying such information alongside
27 Music Publishers’ lyrics in outputs, thereby concealing Anthropic’s infringement from its users,
28 Music Publishers, and other copyright owners. Put differently, Anthropic only wants Claude to

1 train on the expressive content of Music Publishers’ musical compositions, like the lyrics, because
2 that content generates valuable output for Anthropic’s users and customers, while CMI generates
3 no value for Anthropic and, more importantly, provides copyright owners with direct evidence of
4 Anthropic’s infringements.

5 104. Anthropic’s leadership has a history of impermissibly altering copyrighted works
6 to remove CMI in the exact same way. Prior to founding Anthropic, Dr. Amodei helped develop
7 multiple datasets used to train OpenAI’s GPT models, including the “WebText” and “WebText2”
8 datasets. Dr. Amodei and his OpenAI colleagues created the WebText database by “scrap[ing] all
9 outbound links from Reddit, a social media platform,” then “extract[ing] the text” using “a
10 combination of the Dragnet and Newspaper content extractors.”¹⁷ Likewise, Dr. Amodei and
11 several future Anthropic founders and employees—including Mr. Mann, Jared Kaplan, Tom
12 Brown, Jack Clark, and Amanda Askell—developed the WebText2 dataset “by scraping links over
13 a long[] period of time” and extracting their text with the Newspaper extractor.¹⁸

14 105. Today, Anthropic uses the same techniques to remove CMI from Claude’s training
15 corpus with the knowledge that doing so will conceal Anthropic’s infringement of Music
16 Publishers’ works. As early as May 2021, Mr. Mann, Jared Kaplan, and other members of
17 Anthropic’s senior leadership team weighed the merits of using various extraction tools like
18 Newspaper, Readability, and jusText, which are capable of separating copyright notices from the
19 footers of webpages. In June 2021, Mr. Mann and Dr. Kaplan decided not to use jusText because
20 that tool left behind too much “useless junk,” *Concord Music Group, Inc. et al v. Anthropic PBC*,
21 5:24-cv-03811-EKL (N.D. Cal.) (“*Concord I* Docket”), ECF No. 419-1 at ¶ 7, including the critical
22 copyright notices contained in webpage footers—as compared to Readability and Newspaper. *See*
23 *also Concord I* Docket, ECF No. 461 at 2. Indeed, one internal Anthropic chat message revealed
24

25 ¹⁷ Alec Radford et al., *Language Models Are Unsupervised Multitask Learners*, 3, OPENAI
26 (2019), https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

27 ¹⁸ Tom B. Brown et al., *Language Models Are Few-Shot Learners*, 8–9, ARXIV (July 22, 2020),
28 <https://arxiv.org/pdf/2005.14165>; Jared Kaplan et al., *Scaling Laws for Neural Language Models*, 7, ARXIV (Jan. 23, 2020), <https://arxiv.org/pdf/2001.08361>.

1 in prior litigation reflects Mr. Mann’s and Dr. Kaplan’s dissatisfaction with jusText because it failed
2 to remove a copyright owner name and a “© 2019” copyright notice from the footer of a scraped
3 webpage. *Id.* at 8. By contrast, the Anthropic co-founders hailed Newspaper as “a significant
4 improvement” because it reliably removed footers, copyright owner names, and copyright notices.
5 *Id.* Thus, because Newspaper removed CMI more effectively, Anthropic purposefully selected that
6 tool to remove copyright notices and other CMI from Music Publishers’ lyrics and other
7 copyrighted works, which Anthropic continues to do to this day for the purpose of concealing its
8 infringement and making it more difficult to identify copyrighted works within its training corpus.

9 106. Anthropic intentionally applied these same extraction methods to the Common
10 Crawl dataset it downloaded. Specifically, the version of Common Crawl that Anthropic added to
11 Claude’s training corpus contains copies of Music Publishers’ lyrics that were scraped from the
12 websites of Music Publishers’ authorized licensees, including MusixMatch and LyricFind, among
13 others. When MusixMatch, LyricFind, and other licensees display Music Publishers’ lyrics on their
14 respective websites, those lyrics are accompanied by corresponding CMI, per the terms of their
15 licenses with Music Publishers. Anthropic knew when it downloaded Common Crawl that the
16 dataset contained Music Publishers’ copyrighted lyrics and corresponding CMI, including because
17 the company purports to undertake due diligence on all data it uses for training. After copying the
18 contents of Common Crawl, including the copies of Music Publishers’ lyrics therein, Anthropic
19 used extraction tools and algorithms to strip corresponding CMI to curate its own training datasets,
20 including to train its latest Claude models.

21 107. Anthropic also intentionally removes and alters Music Publishers’ CMI when its
22 Claude AI models generate and distribute outputs containing Music Publishers’ lyrics. Anthropic’s
23 executives are well aware that their AI models have a tendency to “memorize” large amounts of
24 data ingested during training. Anthropic’s AI models subsequently and routinely regurgitate this
25 memorized data, including verbatim copies of Music Publishers’ lyrics via output to users
26 unaccompanied by the corresponding copyright notice, song title, songwriter, or other critical
27 identifying information, consistent with Anthropic’s training and design. That is particularly the
28 case when, for example, Claude is prompted to generate purportedly “new” or “original” lyrics in

the style of an existing song or songwriter, and the model generates output copying existing lyrics while failing to provide proper attribution via CMI.

108. Anthropic knowingly distributes outputs containing Music Publishers’ lyrics without corresponding CMI because its objective is to reproduce the lyrics’ expressive content while excluding what it considers “useless junk” or “boilerplate.” *Concord I* Docket, ECF No. 419-1 at ¶ 7. Generating output without CMI further benefits Anthropic because it conceals Anthropic’s infringement from users, Music Publishers, and other copyright owners. Anthropic knows this and removes CMI intentionally for these purposes. Discovery will reveal additional evidence of Claude’s distributing outputs to third-party users without the original CMI, which facts are in the possession and control of Anthropic.

E. Anthropic’s Infringement via AI Model Training

109. As described above, Anthropic claims that training Claude requires an enormous amount of text, which Anthropic compiles into a “training corpus” that it uses to train a “bare LLM model” that will later become Claude. The training corpus that Anthropic uses to build its Claude models contains unauthorized copies of Music Publishers’ copyrighted musical compositions that Anthropic scraped, downloaded, torrented, or copied by other means from the internet, third-party datasets, and pirated sources. Accordingly, each time Anthropic trains a new version of Claude, it violates Music Publishers’ copyrights anew. The training process proceeds as follows.

110. First, after it has assembled and “cleaned” a suitable training corpus, Anthropic “encodes” that training corpus into the computer memory of a “bare LLM model.” This encoding process involves copying, dividing, and converting collected text into units known as “tokens,” which are words or parts of words and punctuation, for storage. Anthropic converts these tokens into numerical “vectors” to help the models recognize each token in their algorithms. Vectors are calculated based on each token’s proximity to others in a given model’s training corpus (*i.e.*, vectors help the model understand a given word’s relation to other specific words, revealing their most frequent meanings and uses). During this encoding process, Anthropic removes any remaining CMI, including that of Music Publishers’ musical compositions.

111. After encoding, Anthropic begins “finetuning” the AI model through a process of

“supervised” and “reinforcement learning.” Supervised and reinforcement learning both rely on human and AI feedback and may involve additional copying of Music Publishers’ works. The goal of supervised learning is to get the model to exhibit a desired response or behavior as a baseline. Once the model exhibits that desired response or behavior, Anthropic employs human contract workers to further prompt and “reward” the model when it continues to exhibit the desired responses and behaviors. Anthropic repeats the reinforcement learning process until it is satisfied with the product, at which point the model is released as a new version of “Claude.”

112. With every new Claude model that it trains, Anthropic restarts its text collection, training, and finetuning processes, which constitute new and independent infringements of Music Publishers’ musical compositions. Anthropic is well-aware that its training data contains unauthorized copies of Music Publishers’ musical compositions and countless other copyrighted works. Anthropic has the ability to exclude Music Publishers’ works from its training corpus (or pay for them) but has chosen to infringe copyrights instead.

113. Beyond its commercial Claude models, Anthropic also creates countless additional copies of Music Publishers’ lyrics when it creates “data mixes” for internal experiments and research models, the learnings from which Anthropic applies in its commercial enterprise. Like most AI companies, Anthropic maintains a repository of all data mixes that its employees and researchers have used and experimented with to train both commercial and non-commercial models. Each new data mix constitutes another copy of Music Publishers’ works contained therein.

114. Anthropic closely guards the specific sources of text that it uses to train Claude and only publishes vague descriptions about the data it uses to train each model. That said, Anthropic has repeatedly admitted that it trains its Claude models on “a proprietary mix of publicly available information on the Internet” and “non-public data from third parties, data provided by data labeling services and paid contractors, and data [it] generate[s] internally.”¹⁹ Most recently, Anthropic

¹⁹ See, e.g., *Claude 3.7 Sonnet System Card*, 3, ANTHROPIC (Feb. 25, 2025), <https://assets.anthropic.com/m/785e231869ea8b3b/original/claude-3-7-sonnet-system-card.pdf>; see also *The Claude 3 Model Family: Opus, Sonnet, Haiku*, 3, ANTHROPIC (last updated Oct. 22, 2024), <https://assets.anthropic.com/m/61e7d27f8c8f5919/original/Claude-3-Model-Card.pdf>; *Model Card and Evaluations for Claude Models*, 2, ANTHROPIC (last updated Nov. 21, 2023), <https://www-cdn.anthropic.com/5c49cc247484cecf107c699baf29250302e5da70/claude-2-model->

1 admitted that it trained its latest models, including Claude Sonnet 5 (June 2026) and Claude Fable
2 5 (June 2026), on “a proprietary mix of publicly available information from the internet, public
3 and private datasets, and synthetic data generated by other models” and that it “use[s] a general-
4 purpose web crawler [called ClaudeBot] to obtain [training] data from public websites.”²⁰
5 Anthropic made similarly sparse disclosures in connection with earlier models.

6 115. Anthropic refuses to disclose the specific sources and composition of its training
7 data because it does not want Music Publishers (and other rightsholders) to know that it is using
8 their copyrighted materials in training without permission. But Claude’s ability to generate
9 verbatim and near-verbatim outputs of Music Publishers’ lyrics in response to user prompts
10 confirms that Anthropic trained Claude on Music Publishers’ copyrighted musical compositions.
11 If not for Anthropic’s copying and use of Music Publishers’ lyrics to train its AI models, Claude
12 would not be capable of reproducing verbatim or near-verbatim copies of Music Publishers’ works
13 as outputs.

14 116. As noted above, Anthropic has admitted that it used data derived from the
15 “Common Crawl” and “The Pile” datasets to train its Claude models. The Common Crawl dataset
16 includes Music Publishers’ copyrighted lyrics that were scraped from the internet, including from
17 websites maintained by Music Publishers’ licensees authorized to display Music Publishers’
18 copyrighted lyrics, such as MusixMatch and LyricFind. Similarly, The Pile includes Music
19 Publishers’ copyrighted lyrics by virtue of the fact that The Pile incorporates numerous pirated
20 books from the predecessor Books3 dataset. Again, Books3 is well known to be comprised of
21 hundreds of thousands of pirated books obtained through torrenting, and includes numerous books
22 containing Music Publishers’ works, such as *VH1’s 100 Greatest Songs of Rock & Roll* (featuring
23 “Maggie May” and other works), and *Alternative Rock Sheet Music Collection* (featuring “We Are
24

25 _____
26 card.pdf.

27 ²⁰ *System Card: Claude Sonnet 5*, ANTHROPIC (June 30, 2026), [https://www-](https://www-cdn.anthropic.com/78073f739564e986ff3e28522761a7a0b4484f84.pdf)
28 [cdn.anthropic.com/78073f739564e986ff3e28522761a7a0b4484f84.pdf](https://www-cdn.anthropic.com/78073f739564e986ff3e28522761a7a0b4484f84.pdf); *System Card: Claude*
Fable 5, ANTHROPIC (June 9, 2026), [https://www-](https://www-cdn.anthropic.com/78073f739564e986ff3e28522761a7a0b4484f84.pdf)
[cdn.anthropic.com/78073f739564e986ff3e28522761a7a0b4484f84.pdf](https://www-cdn.anthropic.com/78073f739564e986ff3e28522761a7a0b4484f84.pdf).

Young” and other works), among others.²¹ Because The Pile also includes the “YouTube Subtitles” dataset, comprised of “all human generated closed captions” for 173,651 YouTube videos, the dataset likewise contains lyrics of Music Publishers’ works from those YouTube videos.”²²

117. Anthropic copied both the Common Crawl and The Pile datasets, among other third-party datasets, and processed their contents for use as AI training data. In processing those datasets, Anthropic intentionally used extraction algorithms designed to remove or alter CMI accompanying Music Publishers’ lyrics, including the copyright notice, copyright owner names, songwriter credits, performing artist name, and song title. The contents of The Pile and the Common Crawl datasets are publicly known, and therefore Anthropic knew or should have known that when it copied those datasets that they included Music Publishers’ copyrighted works.

118. Anthropic trained and still trains its Claude models with information that Anthropic itself scraped and copied from the internet without permission, including copies of Music Publishers’ copyrighted lyrics copied directly from the websites of authorized licensees.

119. Anthropic also used portions of the LibGen and PiLiMi datasets to train its AI models. Anthropic has previously represented that it only used those datasets in one non-commercial research model, *see Bartz Docket*, ECF No. 560-12 at 6, but its later litigation positions belie that representation. For example, in public filings, Anthropic has refused to remove confidentiality designations on its pirated LibGen and PiLiMi datasets, claiming that de-designation “would reveal precisely the contents of the [LibGen or PiLiMi] dataset that Anthropic possessed and used as part of its LLM development.” *Concord I Docket*, ECF No. 450 at 5. Additionally, Anthropic has admitted to using vast amounts of “synthetic data”—i.e., output generated by other Anthropic AI models—as training data for its Claude models. Anthropic has

²¹ See Alex Reisner, *Revealed: The Authors Whose Pirated Books Are Powering Generative AI*, THE ATLANTIC (Aug. 19, 2023), <https://www.theatlantic.com/technology/archive/2023/08/books3-ai-meta-llama-pirated-books/675063/>; *AI Watchdog: Books3*, THE ATLANTIC (Sept. 10, 2025), <https://www.theatlantic.com/technology/archive/2025/09/dataset-books3/683662/>.

²² Leo Gao et al., *The Pile: An 800GB Dataset of Diverse Text for Language Modeling*, 2, ARXIV (Dec. 31, 2020), <https://arxiv.org/pdf/2101.00027>; Alex Reisner, *Search the YouTube Videos Secretly Powering Generative AI*, PROOF (Jul.16, 2024), <https://www.proofnews.org/youtube-ai-search/>.

1 trained its commercial Claude models on at least some synthetic data originating from an AI model
2 that it previously trained with pirated works from LibGen and PiLiMi.

3 **F. Anthropic’s Infringement via AI Outputs**

4 120. Anthropic’s infringement continues post-training in the form of outputs generated
5 by Claude that contain verbatim copies, and which act as market substitutes of Music Publishers’
6 lyrics.

7 **1. Anthropic’s Infringement via AI Outputs Containing Verbatim, Near-**
8 **Verbatim, and Derivative Copies of Music Publishers’ Works**

9 121. Anthropic’s Claude AI models routinely generate verbatim copies of the lyrics to
10 Music Publishers’ musical compositions in response to user prompts. These verbatim outputs are
11 the direct result of, and only possible due to, Anthropic’s unauthorized training of its models with
12 Music Publishers’ lyrics. Each Claude output containing a verbatim copy of Music Publishers’
13 lyrics represents a violation of Music Publishers’ exclusive rights to reproduce and display those
14 lyrics—rights that Music Publishers monetize through a ready commercial licensing market.

15 122. Verbatim output of specific text is a function of AI “memorization,” which is a well-
16 known phenomenon that occurs during the training process. In July 2020, Dr. Amodei and Mr.
17 Mann observed that “[a] major methodological concern with language models pretrained on a
18 broad swath of internet data, particularly large models with the capacity to memorize vast amounts
19 of content, is potential contamination of downstream tasks by having their test or development sets
20 inadvertently seen during pre-training.”²³ Prior litigation reveals that Anthropic employees put it
21 more plainly: “Large LMs memorize A LOT, like a LOT.” *See Concord I* Docket, ECF No. 461 at
22 8 (quoting First Amended Complaint ¶¶ 108–09). Other unsealed litigation filings show Anthropic
23 employees openly discussing how Claude “learn[s] from the training set with a high degree of
24 memorization” and “reconstructs (copyrighted) training data with little prompting effort.” *Bartz*
25 Docket, ECF No. 555-6 at 3. This includes Anthropic’s AI models’ memorizing and regurgitating

26
27 ²³ Tom B. Brown et al., *Language Models Are Few-Shot Learners*, 9, ARXIV (July 22, 2020),
28 <https://arxiv.org/pdf/2005.14165>; Jared Kaplan et al., *Scaling Laws for Neural Language Models*, 7, ARXIV (Jan. 23, 2020), <https://arxiv.org/pdf/2001.08361>.

1 Music Publishers’ lyrics specifically, generating verbatim or near-verbatim copies, and
2 reproducing the “heart” of those works in response to a wide range of user prompts.

3 123. Anthropic designed and trained Claude to generate verbatim copies and derivatives
4 of Music Publishers’ lyrics. During the finetuning process, Anthropic hired teams of temporary
5 workers to “[c]hat with the [Claude] AI to get help with any text-based task.”²⁴ Anthropic
6 communicated daily with these workers over Slack to provide guidance and feedback. Anthropic
7 provided workers with written example tasks that they could assign the model, including
8 “suggesting songs based on your favorite music” or “ask[ing] models to re-write text with style,
9 content, and formatting changes or requests.”²⁵ Consequently, the workers repeatedly prompted
10 the AI models for Music Publishers’ copyrighted lyrics, encouraging the models to generate output
11 containing those lyrics.

12 124. Anthropic’s knowledge and control over the outputs of its AI models is further
13 underscored by its implementation of so-called “guardrails” into its AI models in response to other
14 ongoing lawsuits that exposed Anthropic’s infringing conduct. These “guardrails” purport to
15 restrict Claude from generating verbatim copies of copyrighted works in response to user prompts.
16 But these guardrails are easily circumventable by simply “re-prompting” Claude for specific
17 content, including Music Publishers’ lyrics. Anthropic is aware of these deficiencies and has still
18 failed to remove Music Publishers’ lyrics from its training data or adopt guardrails that prevent the
19 output of Music Publishers’ lyrics in response to re-prompting.

20 125. Notwithstanding Anthropic’s “guardrails,” scientific literature confirms that Claude
21 will still deliver large amounts of copyrighted content as output,²⁶ such as Music Publishers’ lyrics.
22 Discovery will reveal additional evidence of Claude users making lyric- and song-related requests
23 and Claude generating output that copies lyrics in connection with these and other requests, which
24

25 ²⁴ Yuntao Bai et al., *Training a Helpful and Harmless Assistant with Reinforcement Learning*
26 *from Human Feedback*, 63–65, ARXIV (Apr. 12, 2022), <https://arxiv.org/pdf/2204.05862>.

27 ²⁵ *Id.* at 65, 67.

28 ²⁶ See, e.g., Ahmed Ahmed et al., *Extracting books from production language models*, ARXIV
(Jan. 6, 2026), <https://arxiv.org/html/2601.02671v1#bib>.

facts are in the possession and control of Anthropic. Indeed, recently unsealed *Bartz* filings revealed that Anthropic exploits the pirated books it illegally torrented from LibGen in connection with its guardrails, including those containing illegal copies of Music Publishers’ musical compositions. *See Bartz* Docket, ECF No. 554-17 at 93:23–94:14; *see also id.* ECF No. 558 at 17.

126. One dataset that Anthropic used to finetune Claude included numerous requests for lyrics to Music Publishers’ and others’ copyrighted song lyrics, such as, “Do you know the lyrics to ‘I am The Walrus’? . . . Can you give me the first verse?”²⁷ These finetuning records also included various other lyric-related prompts and outputs, such as requests to generate derivatives of Music Publishers’ and others’ lyrics, including, “The following are the lyrics from Eminem’s song [‘Not Afraid’] and I’d like you to change them into Beyonce’s style.” In one such finetuning exchange, Anthropic’s AI model confirmed it has “an extensive database of songs and their lyrics” and that its “repertoire of songs is very extensive and varied,” before outputting lyrics to “Hallelujah” by Leonard Cohen. Subsequently, the Anthropic finetuning reviewer “chose” the AI output copying these lyrics verbatim, and “rejected” output with inaccurate lyrics, thereby encouraging the model to output other song lyrics verbatim in response to similar requests. Other finetuning prompts included similar lyric-related requests for a range of other copyrighted songs, including those belonging to Music Publishers.

127. Anthropic’s own employees also frequently prompted Claude for song lyrics and derivatives of lyrics, including for copyrighted works owned by Music Publishers and others, when developing, testing, and using Claude models.

128. After Claude’s release to the general public, third-party users have made and continue to make similar requests for copies of Music Publishers’ lyrics, and Claude has generated and continues to generate responses reproducing those lyrics, in violation of Music Publishers’ rights. Anthropic knows that Claude’s outputs copy Music Publishers’ lyrics verbatim because it closely monitors and analyzes user interactions. Indeed, Anthropic collects Claude prompts and

²⁷ ANTHROPIC, *HuggingFace, Dataset Card for HH-RLHF*, HUGGINGFACE, <https://huggingface.co/datasets/Anthropic/hh-rlhf> (dataset including quoted Claude finetuning prompts).

1 outputs to study how the model is used before publishing certain of those analyses.²⁸ Anthropic's
2 own analysis of Claude user data reflects user requests to "[h]elp me find, analyze, or modify song
3 lyrics" and to "[t]ranslate songs or lyrics between languages," among many other prompts seeking
4 Music Publishers' and others' lyrics and corresponding output copying those lyrics.

5 129. For example, third-party users have queried Claude for, and received, output
6 containing verbatim copies of Music Publishers' lyrics in violation of Music Publishers' copyrights.
7 Public filings from pending copyright infringement litigation against Anthropic reveal that its
8 Claude AI models have produced infringing output copying verbatim lyrics to works owned in part
9 by Music Publishers, such as "Uptown Funk," "Redbone," "Stay," "Scars to Your Beautiful,"
10 "California Gurls," and "We Belong Together," in response to third-party user requests, as well as
11 infringing output similarly copying lyrics to over 100 works owned by other copyright owners,
12 demonstrating the many different ways in which Claude regurgitates Music Publishers' and others'
13 lyrics in response to a range of such prompts.²⁹ Discovery will reveal additional evidence of
14 Claude users making similar lyric-related requests and Claude generating output copying Music
15 Publishers' lyrics to the works identified in Exhibit B in connection with these and other requests,
16 which facts are in the possession and control of Anthropic.³⁰

17 130. Anthropic's Claude has generated output copying Music Publishers' lyrics in a
18 range of other ways as well. For example, Claude generates purportedly "new" song lyrics that
19 incorporate and copy Music Publishers' song lyrics, or otherwise rely on those lyrics from model
20 training. The result is frequently a "mash-up" of portions from one copyrighted work alongside
21 portions of another. In other instances, when prompted to write poems or lyrics "in the style of"
22 specific artists, Claude generates outputs reflecting the lyrics to Music Publishers' musical
23 compositions. Separately, when Claude is prompted to provide the chord progressions for a given
24

25 ²⁸ See, e.g., Alex Tamkin et al., *Clio: Privacy-Preserving Insights into Real-World AI Use*,
26 ARXIV (Dec. 18, 2024), <https://arxiv.org/pdf/2412.13678>.

27 ²⁹ *Concord I* Docket, ECF No. 728-2 at 238, 719, 724, 894, 913, 939; *id.* ECF No. 595-2 at 28,
164, 244, 253, 290, 306.

28 ³⁰ Music Publishers intend to seek leave to amend this Complaint at an appropriate time to
provide an expanded list of musical compositions infringed by Defendants in this manner.

song, with or without reference to the lyrics, the AI model will often generate output containing Music Publishers' copyrighted lyrics alongside those chords. Claude even copies and distributes Music Publishers' lyrics when it is not asked to do so. Indeed, when Claude is prompted to write a song about a given topic (*i.e.*, without reference to any particular song, artist, or title), it will often respond by generating lyrics that copy directly from portions of Music Publishers' copyrighted lyrics. Claude has generated output copying Music Publishers' lyrics in connection with a myriad of other user requests, which facts are in the possession and control of Anthropic.

131. Anthropic intended for Claude to be able to respond to these and other lyrics-related requests and to reproduce Music Publishers' lyrics in output in this manner. For example, in one early internal record, an Anthropic employee tested Claude's capabilities by prompting it to "[w]rite a coherent poem made up of fragments" of "lyrics from Beatles, Bob Dylan, and other classics from the 60s/70s." *See Concord I* Docket, ECF No. 337 at ¶ 112.

132. Music Publishers have never authorized or licensed Anthropic to copy their copyrighted lyrics or otherwise exploit their lyrics in any of the foregoing manners.

2. Anthropic's Infringement via AI Output Relying on Music Publishers' Works to Create Market Substitutes

133. By including Music Publishers' works in the training data for Claude, Anthropic knowingly enables Claude to not only reproduce verbatim copies and AI-generated derivatives of Music Publishers' lyrics to users upon a user's request, but to also generate market substitutes capable of directly competing with Music Publishers' works. Indeed, every time a user or customer prompts Claude to generate "new" or "original" song lyrics, the model relies, at least in part, on what it has learned about song and lyric construction from the unauthorized copies of Music Publishers' lyrics that Anthropic included in the model's training data.

134. Claude's ability to generate "new" or "original" lyrics is intentional. Anthropic's express goal is to create a human-like intelligence capable of almost all intellectual tasks, including the ability to write song lyrics. Indeed, Anthropic publicly tracks the ways in which more than 1,000 human occupations, including "Lyricists," are impacted or augmented by AI models like

Claude.³¹ One category of request that Anthropic admits it regularly receives from “Lyricists” is a directive for Claude to “write narrative, dramatic, lyric, or other types of poetry.”³²

135. The ability of generative AI models like Claude to write lyrics comes as no surprise to Defendants. Recent studies show that lay readers actually favor “AI-generated text for writing quality,”³³ while others confirm that writing jobs, like those of lyricists, are most vulnerable to being replaced by AI models.³⁴ AI technology “has had an especially swift impact on songwriting” and the United States is now “in the midst of a full-blown A.I. music moment,” with a flood of AI-generated songs topping music charts and streaming platforms.³⁵

136. The U.S. Copyright Office has also observed that “where a generative AI model’s outputs, even if not substantially similar to a specific copyrighted work, compete in the market for that type of work,” including in the case of song lyrics, the outputs may dilute “[r]oyalty pools.”³⁶ This is precisely the kind of harm that Music Publishers face as a result of Claude generating “new” or “original” lyrics after training on Music Publishers’ lyrics. Indeed, Music Publishers generate substantial revenue from digital service providers like Spotify and Apple Music through pro-rata royalty pools that are shared with other rightsholders. Songs that use AI-generated lyrics reduce the total share of streaming revenue to which Music Publishers are entitled because the AI-generated works reduce the total number of streams that would otherwise be attributable to Music

³¹ *Anthropic Economic Index*, ANTHROPIC (last updated June 26, 2026), <https://www.anthropic.com/economic-index>.

³² *Id.*

³³ Tuhin Chakrabarty, Jane C. Ginsburg, & Paramveer Dhillon, *Readers Prefer Outputs of AI Trained on Copyrighted Books over Expert Human Writers*, COLUMBIA PUBLIC LAW RESEARCH PAPER NO. 5606570 (last rev. Aug. 19, 2026), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5606570.

³⁴ Ole Teutloff et al., *Winners and losers of generative AI: Early Evidence of Shifts in Freelancer Demand*, 235 J. OF ECON. BEHAVIOR & ORG. 106845 (2025), at 1.

³⁵ Kyle Chakha, *That New Hit Song on Spotify? It Was Made by A.I.*, THE NEW YORKER (Nov. 12, 2025), <https://www.newyorker.com/culture/infinite-scroll/that-new-hit-song-on-spotify-it-was-made-by-ai>.

³⁶ UNITED STATES COPYRIGHT OFFICE, *Copyright and Artificial Intelligence, Part 3: Generative AI Training*, 65 (May 2025), <https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>.

Publishers' works.

137. Anthropic's users and customers use Claude for the purpose of generating "new" or "original" lyrics to be included in songs that compete with Music Publishers' works for streaming and other revenue. Given the nature of AI-generated material, and the lack of sufficient means to detect such material, information regarding the frequency and degree to which this occurs is within Anthropic's custody and control. Discovery will reveal the full extent to which Anthropic's users and customers rely on Claude to generate market substitutes to Music Publishers' works.

G. Anthropic's Profits from Its Infringement of Music Publishers' Works

138. Anthropic has benefitted immensely from exploiting the copyrighted works of Music Publishers and others in connection with its AI models. At present, Anthropic is regarded as the most valuable AI start-up company in the world, valued at \$965 billion.³⁷ The company's principal investors include blue-chip companies like Amazon, Google, and others.

139. One of the main reasons that Anthropic's AI models are so popular and valuable is because Anthropic has trained those models on a text corpus comprised of copyrighted works, including Music Publishers' lyrics, without compensating rightsholders. As such, Anthropic's astronomical valuation is derived, at least in part, from using Music Publishers' copyrighted lyrics to train Claude.

140. Anthropic's revenue generation has also grown exponentially in recent years. In February 2026, Anthropic reported run-rate revenue of \$14 billion and disclosed that the number of customers spending over \$100,000 annually "has grown 7x in the past year."³⁸

141. Anthropic monetizes and derives substantial value from the Claude models trained on Music Publishers' works by limiting free usage of Claude and encouraging users to sign up for

³⁷ See *supra*, notes 3–4.

³⁸ *Anthropic raises \$30 billion in Series G funding at \$380 billion post-money valuation*, ANTHROPIC (Feb. 12, 2026), <https://www.anthropic.com/news/anthropic-raises-30-billion-series-g-funding-380-billion-post-money-valuation>.

1 “Claude Pro,” which comes with a subscription fee of \$20 per month or \$200 per year, or the “Max
2 20x plan,” which offers “20x more usage than Pro” for \$200 per month.³⁹

3 142. At all relevant times, Anthropic has derived enormous financial benefit from
4 exploiting Music Publishers’ lyrics in training and outputs through the sale of a variety of
5 subscription services.

6 **H. Anthropic’s Infringement Inflicts Serious and Continued Harm**

7 143. Anthropic’s unlawful conduct has caused and continues to cause substantial and
8 irreparable harm to Music Publishers and their songwriters.

9 144. Music Publishers’ musical compositions are creative intellectual property at the
10 core of copyright protection under the Copyright Act. Anthropic is a for-profit commercial
11 enterprise engaged in extensive reproduction and distribution of Music Publishers’ works without
12 authorization from or compensation to Music Publishers.

13 145. Anthropic’s use of BitTorrent to torrent Music Publishers’ works from pirate
14 websites is particularly damaging because each torrented file can instantaneously be distributed
15 hundreds or thousands of times through the “swarm.” Anthropic’s use of BitTorrent to obtain
16 pirated works from known illegal websites sustains and normalizes this infringing protocol and
17 websites, causing and prolonging harm to Music Publishers.

18 146. Anthropic’s removal of CMI from Music Publishers’ copyrighted works impairs
19 Music Publishers’ ability to enforce their copyrights and protect their works from further
20 exploitation by the Claude models and users.

21 147. Anthropic’s use of Music Publishers’ copyrighted works in training and generating
22 outputs devalues the labor of Music Publishers and their songwriters, deprives them of due
23 compensation, and strips control of how their works are used, reproduced, and adapted.

24
25
26 ³⁹ See *How do usage and length limits work?*, ANTHROPIC (July 13, 2026),
27 <https://support.claude.com/en/articles/11647753-how-do-usage-and-length-limits-work>; *What is*
28 *the Pro plan?*, ANTHROPIC (June 10, 2026), [https://support.claude.com/en/articles/8325606-](https://support.claude.com/en/articles/8325606-what-is-the-pro-plan#h_62ccc00135)
[what-is-the-pro-plan#h_62ccc00135](https://support.claude.com/en/articles/8325606-what-is-the-pro-plan#h_62ccc00135); *What is the Max plan?*, ANTHROPIC (Aug. 7, 2026),
https://support.claude.com/en/articles/11049741-what-is-the-max-plan#h_cfd2904008.

1 148. Anthropic has never sought nor obtained any license to lawfully exploit Music
2 Publishers' copyrighted works for any use, let alone for AI training data or output. Anthropic
3 enriches itself through the uncompensated exploitation of Music Publishers' and their songwriters'
4 labor, reducing their incentive to invest in, support, and expand present and future creative efforts.

5 149. Anthropic's unauthorized use harms existing and potential markets for Music
6 Publishers' lyrics, including the market to license AI training data and others, by undermining
7 Music Publishers' ability to control access to their work absent legal action, and by normalizing
8 Anthropic's infringing conduct among potential licensees similarly situated to Anthropic.

9 150. Anthropic also disrupts and dilutes the existing market for Music Publishers' lyrics
10 through the generation of "new" and "original" song lyrics that act as market substitutes, and which
11 are the direct and proximate result of Anthropic's predicate copying of Music Publishers' lyrics in
12 connection with AI training.

13 151. Anthropic could immediately stop infringing Music Publishers' musical
14 compositions and remove all of Music Publishers' lyrics from its training data and central library.
15 Anthropic refuses to do so in order to grow Anthropic's business and reap enormous financial gains
16 from the uncompensated labor of Music Publishers and their songwriters.

17 152. Music Publishers file this lawsuit to end Anthropic's ongoing infringement of their
18 rights and to remedy the significant harm Anthropic has caused and continues to cause.

19 **CLAIMS FOR RELIEF**

20 **Count I—Direct Copyright Infringement by Torrenting (Against All Defendants)**

21 153. Music Publishers re-allege and incorporate by reference each and every allegation
22 set forth in the preceding paragraphs 1 through 152 of this Complaint.

23 154. As detailed above, Defendants unlawfully reproduced and distributed to the public
24 Music Publishers' musical compositions, including the sheet music and lyrics contained therein,
25 without Music Publishers' permission or consent. Specifically, Defendants unlawfully reproduced
26 and distributed to the public Music Publishers' copyrighted musical compositions via torrenting,
27 as listed in Exhibit A. Such activity constitutes direct infringement by Defendants of Music
28

Publishers' registered copyrights and exclusive rights under copyright in violation of the Copyright Act, 17 U.S.C. §§ 106(1), (3) and 501.

155. Music Publishers are the legal or beneficial copyright owners of the musical compositions listed in Exhibit A, which are non-exhaustive, exemplary lists.

156. Each infringement by Defendants in and to Music Publishers' musical compositions constitutes a separate and distinct act of infringement.

157. Defendants' acts of infringement are willful, intentional, and purposeful, in disregard of and with indifference to Music Publishers' rights.

158. As a direct and proximate result of Defendants' infringement of Music Publishers' copyrights and exclusive rights, Music Publishers are entitled to statutory damages, pursuant to 17 U.S.C. § 504(c). Alternatively, at Music Publishers' election, pursuant to 17 U.S.C. § 504(b), Music Publishers shall be entitled to their actual damages and Defendants' profits from infringement that are not taken into account in computing the actual damages, as will be proven at trial.

159. Music Publishers are also entitled to attorneys' fees and costs pursuant to 17 U.S.C. § 505.

**Count II—Contributory Infringement by Torrenting (Against Dario Amodei
and Benjamin Mann)**

160. Music Publishers re-allege and incorporate by reference each and every allegation set forth in the preceding paragraphs 1 through 152 of this Complaint.

161. As detailed above, in June 2021, Mr. Mann used the BitTorrent protocol to unlawfully reproduce and distribute to the public via torrenting millions of pirated books from LibGen including many containing Music Publishers' copyrighted musical compositions. In July 2022, other Anthropic employees used the BitTorrent protocol to unlawfully reproduce and distribute to the public via torrenting millions of pirated books from PiLiMi, including many containing Music Publishers' copyrighted musical compositions. Such activity constitutes direct infringement or an unauthorized act in violation of the Copyright Act, 17 U.S.C. §§ 106(1), (3) and 501.

162. Dr. Amodei expressly directed, approved, controlled, and intentionally induced these infringements by Mr. Mann and other Anthropic employees. Without his direction and approval, this activity would not have occurred. Similarly, Mr. Mann expressly directed, approved, controlled, and intentionally induced other Anthropic employees to engage in the same kind of torrenting through PiLiMi that he had done through LibGen. Without his direction and approval, this activity would not have occurred. Dr. Amodei and Mr. Mann were each fully aware that this activity constituted making illegal copies from well-known pirate websites and further distributing those illegal copies to others. Dr. Amodei and Mr. Mann each knew that they were directing the Anthropic employees working under them to make unlawful copies from known pirated sources and to distribute those unlawful copies further. Accordingly, Dr. Amodei and Mr. Mann are contributorily liable for the infringement of Music Publishers' copyrighted musical compositions via torrenting, including the song lyrics contained therein, in violation of Music Publishers' registered copyrights and exclusive rights under copyright in violation of the Copyright Act, 17 U.S.C. §§ 106(1), (3) and 501.

163. Music Publishers are the legal or beneficial copyright owners to the non-exhaustive, exemplary list of musical compositions listed in Exhibit A.

164. Dr. Amodei and Mr. Mann are contributorily liable for the infringement of the works listed in Exhibit A.

165. Each infringement of Music Publishers' musical compositions constitutes a separate and distinct act of infringement.

166. Dr. Amodei's and Mr. Mann's acts of infringement are willful, intentional, and purposeful, in disregard of and with indifference to Music Publishers' rights.

167. As a direct and proximate result of Dr. Amodei's and Mr. Mann's infringement of Music Publishers' copyrights and exclusive rights, Music Publishers are entitled to statutory damages, pursuant to 17 U.S.C. § 504(c). Alternatively, at Music Publishers' election, pursuant to 17 U.S.C. § 504(b), Music Publishers shall be entitled to their actual damages and Dr. Amodei's and Mr. Mann's profits from infringement that are not taken into account in computing the actual damages, as will be proven at trial.

168. Music Publishers are also entitled to attorneys' fees and costs pursuant to 17 U.S.C. § 505.

Count III—Direct Copyright Infringement (Against Anthropic)

169. Music Publishers re-allege and incorporate by reference each and every allegation set forth in the preceding paragraphs 1 through 152 of this Complaint.

170. As detailed above, Anthropic, without Music Publishers' permission or consent, has unlawfully reproduced, distributed to the public, publicly displayed, and/or prepared derivative works based upon Music Publishers' musical compositions. In particular, Anthropic has unlawfully reproduced, distributed to the public, publicly displayed, and/or prepared derivative works based upon Music Publishers' musical compositions, as listed in Exhibit B, through the following acts:

- a. Scraping unauthorized copies of Music Publishers' lyrics that are displayed on the websites of authorized licensees and elsewhere on the internet;
- b. Downloading unauthorized copies of Music Publishers' lyrics from third-party datasets;
- c. Destructively scanning copies of Music Publishers' lyrics from physical and hard-copy books;
- d. Training public and non-public AI models, including Claude, with unauthorized copies of Music Publishers' lyrics;
- e. Generating verbatim and near-verbatim copies of Music Publishers' lyrics in Claude output;
- f. Generating competitive market substitutes to Music Publishers' lyrics; and
- g. Generating unauthorized derivatives of Music Publishers' lyrics, including derivatives "in the style of" Music Publishers' represented songwriters.

171. The foregoing unauthorized uses are non-exhaustive, ongoing, and constitute direct infringement by Anthropic of Music Publishers' registered copyrights and exclusive rights under copyright in violation of the Copyright Act, 17 U.S.C. §§ 106(1)-(3), (5) and 501.

172. Music Publishers are the legal or beneficial copyright owners of the musical compositions listed in Exhibit B attached hereto, which is a non-exhaustive, exemplary list.

1 173. Each infringement by Anthropic in and to Music Publishers' musical compositions
2 constitutes a separate and distinct act of infringement.

3 174. Anthropic's acts of infringement are willful, intentional, and purposeful, in
4 disregard of and with indifference to Music Publishers' rights.

5 175. As a direct and proximate result of Anthropic's wrongful conduct, which is ongoing,
6 Music Publishers have been and will continue to be substantially and irreparably harmed in an
7 amount not readily capable of determination. Music Publishers have no adequate remedy at law.
8 Unless restrained by this Court, Anthropic will cause further irreparable injury to Music Publishers.
9 Music Publishers are entitled to a permanent injunction prohibiting infringement of Music
10 Publishers' copyrights and exclusive rights under copyright.

11 176. As a direct and proximate result of Defendants' infringement of Music Publishers'
12 copyrights and exclusive rights, Music Publishers are entitled to statutory damages, pursuant to 17
13 U.S.C. § 504(c). Alternatively, at Music Publishers' election, pursuant to 17 U.S.C. § 504(b), Music
14 Publishers shall be entitled to their actual damages and Defendants' profits from infringement that
15 are not taken into account in computing the actual damages, as will be proven at trial.

16 177. Music Publishers are also entitled to attorneys' fees and costs pursuant to 17 U.S.C.
17 § 505.

18 **Count IV—Removal or Alteration of Copyright Management Information**

19 **(Against Anthropic)**

20 178. Music Publishers re-allege and incorporate by reference each and every allegation
21 set forth in the preceding paragraphs 1 through 152 of this Complaint.

22 179. The titles, names, and other identifying information about the authors of Music
23 Publishers' musical compositions, and the name and other identifying information about the
24 copyright owners of those musical compositions, including the song lyrics contained therein,
25 constitute copyright management information under the Copyright Act, 17 U.S.C. § 1202.

26 180. As detailed above, Anthropic intentionally removed or altered CMI from Music
27 Publishers' works, without Music Publishers' authorization, in violation of the Copyright Act, 17
28 U.S.C. § 1202(b)(1), and/or distributed Music Publishers' works or copies of Music Publishers'

1 works knowing that CMI has been removed or altered, without Music Publishers’ authorization,
2 in violation of the Copyright Act, 17 U.S.C. § 1202(b)(3). Further, Anthropic has done so knowing
3 or having reasonable grounds to know that it will induce, enable, facilitate, or conceal an
4 infringement of Music Publishers’ works, including concealing Anthropic’s own infringement.

5 181. Anthropic intentionally removed and altered CMI in the process of training
6 Anthropic’s AI models, including by copying Music Publishers’ lyrics from websites containing
7 those lyrics, including the websites of Music Publishers’ licensees, and by applying algorithms
8 known to remove copyright notices and other CMI from that copied text. Anthropic uses third-
9 party training datasets from which it knows such CMI has been removed or from which it
10 intentionally removes CMI. Anthropic does so purposefully so that it can train its AI models on
11 the expressive content of Music Publishers’ copyrighted lyrics in order to reproduce that expressive
12 content. Anthropic intentionally removes and alters corresponding CMI, which it considers
13 “useless junk” and “boilerplate,” at least in part, for the purpose and objective of concealing its
14 infringement. Anthropic also intentionally removed and altered CMI in the output of those AI
15 models, including by generating text that copies Music Publishers’ lyrics verbatim while omitting
16 the CMI that accompanies authorized versions of these works, and in the process of torrenting
17 Music Publishers’ works from pirate libraries.

18 182. Anthropic also distributed Music Publishers’ copyrighted works and copies of those
19 copyrighted works knowing that CMI was removed or altered, including by copying training
20 datasets that contain Music Publishers’ copyrighted lyrics and to which algorithms known to
21 remove CMI had been applied, and knowingly using copies of Music Publishers’ lyrics from which
22 CMI had been removed to train its models, while knowing that those models would “memorize”
23 and regurgitate copies of those works without their CMI in the models’ outputs, as well as in the
24 process of torrenting Music Publishers’ copyrighted works from pirate libraries.

25 183. Anthropic knew or had reasonable grounds to know that these acts would induce,
26 enable, facilitate, or conceal infringement of Music Publishers’ copyrighted works, including by
27 concealing Anthropic’s own infringement of Music Publishers’ lyrics as the input and output of its
28 AI models; concealing other infringements in connection with the unlawful inclusion of Music

Publishers' works in third-party training datasets; and by inducing, enabling, facilitating, and concealing infringements by Anthropic's users, who are not informed that the outputs they receive from Anthropic's AI models contain copyrighted content.

184. Music Publishers are the legal or beneficial copyright owners of the musical compositions listed in Exhibit B attached hereto, which is a non-exhaustive, exemplary list.

185. Anthropic is liable for the removal or alteration of CMI for the works listed in Exhibit B.

186. Anthropic's infringements are willful, intentional, and purposeful, in disregard of and with indifference to Music Publishers' rights.

187. As a direct and proximate result of Anthropic's wrongful conduct, which remains ongoing, Music Publishers have been, and will continue to be, substantially and irreparably harmed in an amount not readily capable of determination. Music Publishers have no adequate remedy at law. Unless restrained by this Court, Anthropic will cause further irreparable injury to Music Publishers. Music Publishers are entitled to a permanent injunction prohibiting Anthropic's unlawful activity.

188. As a direct and proximate result of Anthropic's unlawful activity, Music Publishers are entitled to statutory damages, pursuant to 17 U.S.C. § 1203(c)(3)(B). Alternatively, at Music Publishers' election, pursuant to 17 U.S.C. § 1203(c)(2), Music Publishers shall be entitled to their actual damages and Anthropic's profits from infringement that are not taken into account in computing the actual damages, as will be proven at trial.

189. Music Publishers are also entitled to attorneys' fees and costs pursuant to 17 U.S.C. § 1203(b)(4)-(5).

PRAYER FOR RELIEF

By reason of the facts and circumstances alleged above, Music Publishers seek relief against Anthropic as follows:

a. Judgment on each of the foregoing claims, including that Anthropic directly infringed Music Publishers' copyrights under the Copyright Act; that Dario Amodei and Benjamin

1 Mann directly and contributorily infringed Music Publishers' copyrights under the Copyright Act;
2 that Anthropic unlawfully removed and/or altered Music Publishers' CMI in violation of the
3 Copyright Act; and that such infringements and unlawful activities were willful;

4 b. An order requiring Defendants to pay Music Publishers statutory damages in an
5 amount up to the maximum provided by law, arising from Defendants' willful violations of Music
6 Publishers' rights under the Copyright Act, including in an amount up to \$150,000 per work
7 infringed, pursuant to 17 U.S.C. § 504(c); or in the alternative, at Music Publishers' election, Music
8 Publishers' actual damages and Defendants' profits from the infringement, in an amount to be
9 proven at trial, pursuant to 17 U.S.C. § 504(b);

10 c. An order requiring Anthropic to pay Music Publishers statutory damages in an
11 amount up to the maximum provided by law, arising from Anthropic's removal and/or alteration
12 of Music Publishers' CMI in violation of the Copyright Act, including in an amount up to \$25,000
13 per violation, pursuant to 17 U.S.C. § 1203(c)(3)(B); or in the alternative, at Music Publishers'
14 election, Music Publishers' actual damages and Anthropic's profits, in an amount to be proven at
15 trial, pursuant to 17 U.S.C. § 1203(c)(2);

16 d. An order for such equitable relief under Titles 17, 28, and/or the Court's inherent
17 authority as is necessary to prevent or restrain infringement of Music Publishers' copyrights,
18 including a permanent injunction requiring that Anthropic and its officers, agents, servants,
19 employees, attorneys, directors, successors, assigns, licensees, and all others in active concert or
20 participation with any of them, including Dario Amodei and Benjamin Mann, cease directly or
21 indirectly infringing, or causing, aiding, enabling, facilitating, encouraging, promoting, inducing,
22 or materially contributing to or participating in the infringement of any of Music Publishers'
23 exclusive rights under copyright, including without limitation in the musical compositions in
24 Exhibits A and B;

25 e. An order requiring Defendants to provide an accounting of the training data,
26 training methods, and known capabilities of Anthropic's AI models, including requiring that
27 Anthropic identify the Music Publishers' lyrics and other copyrighted works on which it has trained
28 its AI models, and disclose the methods by which Anthropic has collected, copied, processed, and

1 encoded this training data (including any third parties it has engaged to collect or license such
2 data).

3 f. An order requiring that Defendants destroy under the Court's supervision all
4 infringing copies of Music Publishers' copyrighted works in Defendants' possession or control,
5 and then file a sworn report setting forth in detail the manner in which it has complied with the
6 aforesaid order, pursuant to 17 U.S.C. § 503(b);

7 g. Music Publishers' reasonable attorneys' fees and costs in this action, pursuant to
8 17 U.S.C. § 505 and/or 17 U.S.C. § 1203(b)(4)-(5);

9 h. Pre-judgment and post-judgment interest at the applicable rate on any monetary
10 award made part of the judgment against Defendants; and

11 i. Such other and further relief as the Court deems proper.

12 **JURY TRIAL DEMAND**

13 Music Publishers hereby request a trial by jury.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

1 Dated: August 28, 2026

Respectfully submitted,

2 /s/Benjamin S. Akley

3 **OPPENHEIM + ZEBRAK, LLP**

Matthew J. Oppenheim

Nicholas C. Hailey

4 Corey Miller

(pro hac vice applications to follow)

5 4530 Wisconsin Ave., NW, 5th Floor

Washington, DC 20016

6 Telephone: (202) 480-2999

matt@oandzlaw.com

7 nick@oandzlaw.com

corey@oandzlaw.com

8 Jennifer Pariser

Carly K. Rothman

9 Bret Matera

(pro hac vice applications to follow)

10 461 5th Avenue, 19th Floor

New York, NY 10017

11 Telephone: (212) 951-1156

jpariser@oandzlaw.com

12 carly@oandzlaw.com

bmatera@oandzlaw.com

13 **PRYOR CASHMAN LLP**

14 Benjamin S. Akley (SBN 278506)

1901 Avenue of the Stars, Suite 900

15 Los Angeles, California 90067

Telephone: (310) 683-6641

16 bakley@pryorcashman.com

17 *Attorneys for Plaintiffs*